

DOCUMENT RESUME

ED 154 951

95

RC 010 481

AUTHOR Coulson, John E.
TITLE Overview of the National Evaluation of the Emergency School Aid Act.
INSTITUTION System Development Corp., Santa Monica, Calif.
SPONS AGENCY Bureau of Elementary and Secondary Education (DHEW/OE), Washington, D.C. Div. of Equal Educational Opportunities.
REPORT NO SDC-TM-5236/022/00
PUB DATE Jul 77
CONTRACT OEC-0-73-0831; OEC-0-73-6336
NOTE 59p.

EDRS PRICE MF-\$0.83 HC-\$3.50 Plus Postage.
DESCRIPTORS *Academic Achievement; Black Students; Caucasian Students; *Compensatory Education Programs; Educational Finance; Educational Legislation; Elementary Secondary Education; Expenditure Per Student; *Program Effectiveness; *Program Evaluation; *Resource Allocations; School Demography; School Funds; School Services; Socioeconomic Status; Spanish Speaking; Student Characteristics

IDENTIFIERS *Education Amendments 1975 Title VII

ABSTRACT

An overall evaluation of the Basic and Pilot Grants Programs, two closely related programs authorized under the Emergency School Aid Act (ESAA), and an in-depth study designed to identify specific program approaches associated with project success were conducted. Data were collected over a period of three school years: 1973-74, 1974-75, and 1975-76. The overall evaluation involved national samples of schools with Basic and Pilot projects; and the in-depth study involved a subset of those projects. For the overall evaluation, comparable student, program, and outcome data in both ESAA-funded and non-ESAA schools were collected along with data on district, school, and program characteristics, student achievement, and background characteristics of students and school personnel. The in-depth study, conducted during the second and third years of the overall evaluation, involved intensive classroom observations and staff interviews. This report provides an overview of the evaluation design and results for all three years, emphasizing the similarities and differences in the findings across years. Findings included: all three years, dollars were targeted toward needy school districts, schools, and students, and translated into services directed toward students with needs for such services; strength of administrative leadership was a major factor in the schools' ability to improve student achievement; in general, achievement gains tended to be larger in schools with greater resource availability and utilization. (NQ)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED154951

Overview of the National Evaluation of the Emergency School Aid Act





OVERVIEW OF THE NATIONAL EVALUATION
OF THE EMERGENCY SCHOOL AID ACT

John E. Coulson

July 1977

Emergency School Aid Act National Evaluation
United States Office of Education Contract Numbers
OEC-0-73-0831 and OEC-0-73-6336

System Development Corporation

TM-5236/022/00

MEMBERS OF THE ESAA EVALUATION TEAM

Project Director

John E. Coulson
Diana D. Dingler, Secretary

Planning and Control

Thomas F. Collins, Senior Control Officer
Marilyn Stevenson, Staff Administrator
Susan E. Shehtanian, Secretary

Operations

Raymond B. Stewart, Manager
Jean B. Wellisch, Special Studies Coordinator
Ronald A. Carriere
Gary A. Duck
Anne H. MacQueen
Marilyn P. Ridley, Secretary
Dean R. Lee, Data Collection Coordinator
Emily H. Allison
Cleone Geddes
Frances Jacobson, Secretary

Analysis

Dan G. Ozenne, Manager
Raymond W. Baker
Clarence Bradford
Deborah Brown
William J. Doherty
Sally D. Hanes
Judith A. Hemenway
Richard J. Kwan
Suzanne Navarrete
Eric Nelson-Lund

Documentation

Kean Mantius
Una Vere Katter
Marcia Simonson, Repro Typist

Project Officer, U.S. Office of Education

David Shoemaker (succeeded M. J. Wargo in 1976)

Consultant

Harold Nichols

Subcontractors

American College Testing Program
Center for the Study of Evaluation

ESAA Evaluation National Advisory Panel

Louis Alvarez
Boyd Bosma
William Coffman
Nancy S. Cole
Silas Craft (succeeded Buell Gallagher)
Jess Elliott
Edgar Epps
Seymour Feshbach
Gene V. Glass
Patricia Gurin
Chester W. Harris
Robert D. Hess
Robert Hill
Sam Husk
Lemuel Ignacio
George Jackson
Ruth King
Melvin K. Novick
Thomas F. Pettigrew
Michael Scriven
Daniel L. Stufflebeam
Charles L. Thomas
Charles W. Townsel
Ralph W. Tyler
Tony Vasquez
Meyer Weinberg
David E. Wiley

Former Project Members

Patricia Acosta
Lillian Barros
Aida Bell
Alice Bloch
Alan Cohen
Marilyn Crilley
Janet Chryssanthi
Constance Dolphus

Paula Geiber
Janette Hatch
Mildred Henderson
Michael Hirsch
Earl Houseman, Consultant
James Hoyt
Maron Jones
Ward Keesling, Consultant
Alfred Marcus

Catherine Maurer
Ralph Melaragno
Patrice Metro
Helen Nero
Thomas Ramirez
Leo Reeder, Consultant
Candy Smith
Hunter Temple, Consultant
Nancy Van Gelder

ACKNOWLEDGMENTS

As indicated by the extensive list of project team members shown on the preceding page, this report is based on the work of many individuals. Also gratefully acknowledged, but not listed because of space limitations, are the many helpful administrators and staff in participating school districts, without whose cooperation this study could not have been conducted.

Much of the information in this overview document is summarized from earlier reports, and special thanks are given to the authors of those reports. In particular, Jean Wellisch, Ronald Carriere, Anne MacQueen, Gary Duck, and Al Marcus were responsible for In-Depth Study documentation that served as a major source for certain portions of this report. Other sections are based in part on earlier materials contributed to by Dan Ozenne, Sally Hanes, Clarence Bradford, William Doherty, Gary Duck, Judy Hemenway, and Nancy Van Gelder.

David Shoemaker, the U.S. Office of Education Project Officer for this study, has been highly supportive in the preparation of this overview, and has made several valuable suggestions for improving its interpretability and utility. More importantly, perhaps, both he and his predecessor, Michael Wargo, have been extremely helpful during the course of the study in formulating evaluation goals and in identifying some of the most critical policy issues to be addressed. We are also grateful to George Rhodes, of the Equal Educational Opportunity Program administration, for making information on ESAA awards available to us.

Throughout this evaluation, an extremely important role was played by the ESAA Evaluation National Advisory Panel, whose members are listed on the preceding page of this report. This panel, which included many national leaders in research methodology and in policy issues related to desegregation and social action programs, has made major substantive suggestions as well as helping to provide broad perspective in interpretations of the study's findings. Also

helpful were the suggestions of the Presidentially appointed National Advisory Council on Equality of Educational Opportunity, which reviewed some of the earlier evaluation documentation.

Finally, we wish to thank Kean Martius for her interest and ability in editing this report, and Marcia Simonson and Dee Dingler for their valuable typing support.

This report was prepared by System Development Corporation under United States Office of Education Contract Numbers OEC-0-73-0831 and OEC-0-73-6336. Despite the many important contributions of individuals and groups acknowledged above, the final contents of this report are solely the responsibility of its author and of System Development Corporation. Views expressed in this document are not necessarily the views of the U.S. Office of Education, nor do we imply any endorsement of the report by individuals or groups outside SDC who may be listed above.

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION	1
ESAA Programs Evaluated	1
Major Issues Addressed	2
Design of Evaluation	3
Overall Evaluation Design	3
In-Depth Study Design	6
II. CHARACTERISTICS OF SAMPLE DISTRICTS AND SCHOOLS	9
District Demographic Characteristics	9
District Desegregation Status	9
School Socioeconomic Status	10
School Racial/Ethnic Composition	11
III. STUDENT CHARACTERISTICS AND NEEDS	13
Proportions of Minority and Non-Minority Students	13
Socioeconomic Status of Parents	14
Student Achievement Test Levels	14
IV. SCHOOL FUNDING PATTERNS	17
Procedures for Estimating Per-Pupil Expenditures	17
Results	18
Comparison of Expenditures at Elementary and Secondary Levels	18
Comparison of Expenditures in Treatment and Control Schools	19
V. RESOURCE ALLOCATIONS AND SERVICES DELIVERED	21
Allocations of Supplemental Funds to Different Activities	21
Comparison of Allocations in Treatment and Control Schools	23
Targeting of Resources to Meet Needs	23

TABLE OF CONTENTS (Cont'd)

	<u>Page</u>
VI. ESAA PROGRAM IMPACT ON STUDENT ACHIEVEMENT	25
Cross-Sectional (One-Year) Results	25
Possible Explanations for Year Three Impact Findings	28
Longitudinal (Three-Year) Results	29
VII. RELATIONSHIPS AMONG PROGRAM VARIABLES, STUDENT ACHIEVEMENT, AND SCHOOL CLIMATE	31
Analytic Methodology	31
Results from Overall Evaluation	32
Funding and Allocation Levels	32
Period of Exposure to Instruction	33
Pupil/Teacher Ratio	33
Program Activities Supporting Desegregation	33
Teacher Training	34
Instructional Practices	34
Results from In-Depth Study	35
Organizational Climate	35
Reading and Mathematics Instructional Practices	36
Resource Utilization	38
Equality of Educational Opportunity	38
Parent and Community Involvement	39
VIII. SUMMARY OF FINDINGS AND RELATIONSHIPS TO EARLIER STUDIES	41
Impact Results	41
Relative Effectiveness of Different Program Approaches	42
XI. THE EVALUATION DESIGN, IN HINDSIGHT	49

TABLE OF CONTENTS (Cont'd)

	<u>Page</u>
EARLIER ESAA REPORTS	51
REFERENCES	53

OVERVIEW OF THE NATIONAL EVALUATION
OF THE EMERGENCY SCHOOL AID ACT

I. INTRODUCTION

In conjunction with and under contracts from the U.S. Office of Education, System Development Corporation (SDC) conducted an evaluation of two closely related programs authorized under the Emergency School Aid Act (ESAA)--the Basic and Pilot Grants Programs. The combined Basic/Pilot Program evaluation involved the collection of data over a period of three school years: 1973-74, 1974-75, and 1975-76. The present report provides an overview of the evaluation design and results for all three years, with particular emphasis on similarities and differences in the findings across years. Earlier and more detailed reports on each year's evaluation results included the following:*

The First Year of Emergency School Aid Act (ESAA) Implementation.
System Development Corporation, TM-5236/008/00, September 1975.

The Second Year of Emergency School Aid Act (ESAA) Implementation.
System Development Corporation, TM-5236/009/00, July 1976.

The Third Year of Emergency School Aid Act (ESAA) Implementation.
System Development Corporation, TM-5236/014/00, March 1977.

An In-Depth Study of Emergency School Aid Act (ESAA) Schools:
1974-1975. System Development Corporation, TM-5236/010/00,
July 1976.

An In-Depth Study of Emergency School Aid Act (ESAA) Schools:
1975-1976. System Development Corporation, TM-5236/013/00,
March 1977.

ESAA PROGRAMS EVALUATED

The Basic Program is a desegregation assistance program in which grants are awarded to eligible school districts to encourage the reduction of minority-group isolation, to meet the special needs incident to the elimination of seg-

*A more complete list of relevant ESAA evaluation documentation is given at the end of this report.

regation and discrimination, and to assist elementary and secondary school children in overcoming the educational disadvantages associated with minority-group isolation. Basic Grants may be awarded to any Local Education Agency (LEA) that (a) is implementing a desegregation plan or has adopted and will implement such a plan if assistance is made available; or (b) has a plan to enroll non-resident children in its schools to reduce minority-group isolation; or (c) has no desegregation plan but has minority-group student enrollment exceeding 50 percent, provided that the LEA establishes or maintains at least one integrated school.

In contrast to Basic Grants, Pilot Project Grants are awarded to LEAs for unusually promising projects designed to overcome the adverse effects of minority-group isolation by improving the academic achievement of children in minority-isolated schools (i.e., schools with 50 percent or greater minority enrollment). To be eligible for a Pilot Grant, an LEA must be implementing a desegregation plan or a plan to reduce minority-group isolation that would make it eligible for a Basic Grant. In addition, at least 15,000 minority-group students must be enrolled in the schools of the LEA, or minority-group students must constitute more than 50 percent of the total LEA enrollment.

MAJOR ISSUES ADDRESSED

Although the ESAA evaluation had a wide variety of goals, its major emphasis was on the following six questions:

- What kinds of districts and schools receive support under the ESAA Basic and Pilot Programs?
- What kinds of students participate in these programs?
- What are the funding patterns in ESAA districts and schools?
- How are ESAA expenditures and other funds translated into educational services? How successfully are services targeted toward needy students?

- What is the short-term and long-term national impact of ESAA in terms of Congressionally authorized program goals?
- What kinds of school programs have most successfully produced gains in student outcomes?

DESIGN OF EVALUATION

The ESAA evaluation project had two major components: a large-scale impact study involving national samples of schools with Basic and Pilot projects, and a more intensive, in-depth study involving a subset of those projects, designed to identify specific program approaches associated with project success. For convenience, those two components are referred to in this report as the "overall evaluation" and the "in-depth study."

Overall Evaluation Design

The design of the overall evaluation involved the collection of comparable student, program, and outcome data in both treatment (ESAA-funded) and control (non-ESAA) schools. Achievement tests and school climate instruments* were administered to sample students near the beginning and end of the school year. In addition, questionnaires were administered to superintendents, ESAA coordinators, district business managers, principals, teachers, and students in the spring to obtain data on district, school, and program characteristics, and on background characteristics of students and school personnel.

*The school climate instruments focused on students' perception of racial climate in their schools, and were originally intended to complement the achievement scores as outcome measures for the evaluation. However, the climate instruments were not satisfactory for this purpose, and are not further discussed in this overview, as they produced scores with extremely limited variability over time and across samples. While the reported perceptions were generally positive, this trend as well as the lack of variability may reflect to a considerable degree a tendency for students to select socially desirable responses.

To select the treatment and control schools in Year One (1973-74), pairs of similar, ESAA-eligible schools were identified in sample ESAA districts; one member of each selected pair was randomly assigned to the treatment (ESAA-participating) condition and the other to the control (non-ESAA-participating) condition. Sample students were then drawn randomly across classes in each sample school in grades 3, 4, and 5 in the elementary schools (Basic and Pilot samples) and grades 10, 11, and 12 in the secondary schools (for the Basic sample only, as there were not enough secondary-level Pilot projects to be evaluated). This evaluation design made it possible to compare treatment and control groups that were initially matched in all important characteristics examined except those associated with the award of ESAA funds to the treatment schools. We could then determine whether the two sets of schools--one receiving ESAA funds and the other not receiving such funds--showed corresponding differences in resource allocation, in services offered to the students, and in educational approaches. Only if such program differences were found, would we expect differences in outcome (e.g., in student achievement gains) for the treatment and control schools.

In sampling for Year Two (1974-75) of the overall evaluation, several new districts and schools were added to increase the number of matched pairs, particularly in the Pilot sample and at the secondary level in the Basic sample. At the same time, several districts were dropped from the sample because they did not apply for or were not awarded new grants for 1974-75. All students who had progressed beyond the grade range of concern in the study (i.e., beyond the fifth or twelfth grade) were automatically dropped from the sample. New samples of students were drawn at the third and tenth grades, and students were added as needed at grades 4, 5, 11, and 12 to offset attrition.

In Year Three (1975-76) no new districts or schools were added, as the major emphasis was on retaining as large a longitudinal sample as possible. Again, however, new third- and tenth-grade student samples were drawn, and students were added at other grade levels to offset attrition. Several more schools

were lost to the sample because of discontinued ESAA funding or because of changes in the grade levels participating in ESAA activities. As in the previous year, students who had progressed beyond grade 5 or grade 10 were automatically dropped from the study.

Table 1, below, gives an overview of changes in sample size over the three years of data collection. It shows each year's cross-sectional (one-year) samples and the longitudinal sample (districts and schools that were in the study all three years). Cell entries include only the schools used for impact analyses (i.e., schools in matched treatment-control pairs); samples in Years One and Two included some non-paired treatment schools that were used to augment sample size for descriptive purposes, but that were not used in the impact analyses.

Table 1. Sample Changes Over Years

Sample	Schools in Treatment-Control Pairs			
	Year One	Year Two	Year Three	Longitudinal
Basic Elementary	74	68	48	34
Basic Secondary	24	40	28	16
Pilot Elementary	36	40	30	16

At the student level, the total sample sizes for the cross-sectional impact analyses were 19,573 in Year One; 17,297 in Year Two; and 9,541 in Year Three. The longitudinal sample contained only those students who were in the Year One sample at grade 3 or grade 10 and who were still in the study for Year Three posttesting at the fifth- or twelfth-grade level; that longitudinal cohort included a total of 1,383 students.

Despite the variations in sample size across years, and the sizable attrition rate for the three-year longitudinal sample, there was little evidence of any important systematic bias in the district-level or school-level sample changes. The districts and schools comprising the Year Two and Year Three cross-sectional samples and the longitudinal sample were generally similar in the demographic dimensions examined (e.g., racial/ethnic mix, enrollment sizes, desegregation status, percentages of disadvantaged students) to those in the original Year One sample. One exception was that Basic Secondary districts in the longitudinal sample had higher percentages of minority students than did districts in each year's cross-sectional samples; even here, however, the school samples were very similar.

Student-level analyses of the three years' cross-sectional samples also revealed very few statistically significant changes in sample characteristics, despite the fact that each year's sample tended to have a slightly higher percentage of White students and slightly higher achievement pretest scores than the preceding year. Although the Basic longitudinal sample had a substantially lower percentage of White students than did any of the Basic cross-sectional samples, pretest scores for the cross-sectional and longitudinal samples did not differ markedly. Overall, it appears that most of the relationships described in subsequent sections of this report, particularly with respect to students' achievement growth patterns, probably reflect actual trends rather than artifacts of inter-year sampling variations or of selective attrition.

In-Depth Study Design

The in-depth study was conducted during the second and third years of the overall evaluation (school years 1974-75 and 1975-76). In both years, the major goal was to identify the program and contextual characteristics that differentiated most clearly and consistently between successful and unsuccessful elementary schools. For this purpose, intensive classroom observations and staff interviews were conducted in subsamples of elementary schools (24 schools in 1974-75 and 26 schools in 1975-76) selected from the overall evaluation sample on the basis

of their wide variations in program effectiveness (i.e., in the schools' ability to increase students' reading and math achievement).

Both years of the in-depth study focused upon organizational climate in the school, reading and mathematics instructional practices, resources used in reading and mathematics instruction, and school practices designed to ensure equality of educational opportunity for minority and non-minority students. The first year's design was exploratory in nature; many variables were examined in an effort to identify those that were most strongly related to program success. The second year's effort was confirmatory; it attempted to replicate the first year's findings and to investigate possible underlying causal relationships.

II. CHARACTERISTICS OF SAMPLE DISTRICTS AND SCHOOLS

DISTRICT DEMOGRAPHIC CHARACTERISTICS

In all three evaluation years, the heaviest concentrations of sample districts for the overall evaluation were in the southern portion of the country, with between 55 percent and 70 percent of the districts falling in HEW Regions IV (Southeast) and VI (South Central); this corresponded fairly closely to the distribution of the ESAA award universes, as approximately two-thirds of all districts receiving ESAA grants were in those two regions. No other region had especially large percentages of districts over the three years, although Region III (Mid-Atlantic) constituted around one-tenth of the award universe and sample each year.

ESAA sample districts tended to be urban, with around three-fourths of the districts located in a medium-sized or large city (50,000 population or greater), 15 percent in smaller towns or suburbs, and only around 10 percent in rural communities. The Pilot sample was slightly less urbanized, with about half the districts in medium-to-large cities. Reflecting their urban nature, most sample Basic districts were fairly large; the typical Basic Elementary sample district had an enrollment of around 14,000 students, and Basic Secondary sample districts averaged around 30,000 students. Pilot Elementary districts were somewhat smaller, with an average (median) enrollment of around 8,000. Overall, sample districts were substantially larger, on the average, than those in the ESAA award universes from which the samples were drawn. This size discrepancy results from the sampling strategy's emphasis on selecting districts with two or more pairable ESAA-eligible schools, so that matched treatment and control groups could be formed; such a requirement tends to screen out the smaller districts, particularly at the secondary level.

DISTRICT DESEGREGATION STATUS

Most of the sample districts had achieved a substantial degree of desegregation prior to the initial award of ESAA grants in 1972. That is, the racial/ethnic

7.8 blank

mix of students in each school tended to match fairly well the racial mix averaged across the entire district.* Furthermore, at least with regard to Black and White students, the degree of racial balance in the districts receiving ESAA grants was clearly superior to that in a survey of almost half the nation's districts, conducted by the Office of Civil Rights in 1972. This is a highly positive finding in one respect. It indicates that, quite appropriately, ESAA grants were typically awarded to districts that had already desegregated to a considerable extent, and were presumably used to help the districts meet needs incident to the desegregation. However, this finding also means that there was little room for further improvement in racial balance after the grants were awarded (although a few districts did show reduction in minority isolation during the evaluation period). Thus, reduction in minority isolation did not prove a useful outcome measure for assessing ESAA program impact.

SCHOOL SOCIOECONOMIC STATUS

Sample schools, particularly at the elementary level, had large numbers of economically disadvantaged students, as indicated by the percentages of students sufficiently disadvantaged to be eligible for Elementary and Secondary Education Act (ESEA) Title I services. These percentages varied somewhat across years, but in general they were highest for the Pilot Elementary sample schools (around 60 percent), somewhat lower for Basic Elementary schools (about 45 percent), and considerably lower for the Basic Secondary sample (around 25 percent). The percentages of Title I-eligible students in sample schools were consistently higher than those for the total districts from which those schools were drawn; although the criteria for defining Title I eligibility are not identical, the differences between district-level and school-level percentages strongly suggest that ESAA funds within sample districts were targeted toward some of the neediest (most disadvantaged) schools.

*It is important to note that the desegregation indices used in this study simply reflect, at a fairly gross level, the district/school racial balance. They do not provide a picture of the status of school integration in terms of interpersonal interactions or attitudes.

SCHOOL RACIAL/ETHNIC COMPOSITION

The percentages of minority students enrolled in sample schools differed substantially for the different evaluation samples; they also changed somewhat from one year to the next, but this probably resulted more from the fact that the sample was composed of a somewhat different set of schools than from composition shifts within the schools. Specifically, the percentage of minority students in Basic Elementary schools increased from 44 percent in Year One to 47 percent in Year Three, while that for Basic Secondary schools decreased in the same period from 44 percent to 36 percent. Not surprisingly (since Pilot schools were required under the grant eligibility guidelines to have at least 50 percent minority students), the largest proportions of minority students were in the Pilot sample schools--78 percent in Year One, rising to 85 percent in Year Three.

Minority representation among staffs in the sample schools was consistently lower than among the students. The largest percentage deviation was in the Pilot Elementary sample where, in Year Two, for example, around 66 percent of the students were Black, but only 40 percent of the schools' staff members were Black.

III. STUDENT CHARACTERISTICS AND NEEDS

PROPORTIONS OF MINORITY AND NON-MINORITY STUDENTS

As shown below in Table 2, there was a small increase each year in the percentage of White students for the two elementary-level samples, but, overall, the proportions of minority and non-minority students in the cross-sectional samples did not change markedly over the three years.

Table 2. Changes in Percentages of Minority and Non-Minority Students

Evaluation Sample	Racial/Ethnic Group	Year One	Year Two	Year Three
Basic Elementary	Minority	55	53	49
	Non-minority	45	47	51
Basic Secondary	Minority	46	41	42
	Non-minority	54	59	58
Pilot Elementary	Minority	88	86	85
	Non-minority	12	14	15

In the Pilot Elementary samples, well over three-fourths of all sample students were minority; around two-thirds were Black, one-fifth were Spanish-background, and fewer than 15 percent were White. Basic Elementary sample schools were almost evenly divided between minority and White students, with Black students constituting about 42 percent and Spanish-background students about 8 percent of the total samples. The Basic Secondary samples had the smallest proportion of minority students, with about 40 percent Black, 3 percent Spanish-background, and 57 percent White students.

SOCIOECONOMIC STATUS OF PARENTS

According to the reports of teachers, fewer than half the parents of sample students had completed high school, and fewer than one-fourth of the fathers had completed college. Furthermore, a large percentage of parents were thought by the teachers to have unskilled, low-paying jobs. Although the accuracy of the teachers' information may be questioned, it appears reasonably certain that the sample students were economically as well as educationally disadvantaged. On the average, lower socioeconomic levels were reported for students in minority-isolated (Pilot) schools, and for minority students regardless of their school's desegregation status.

STUDENT ACHIEVEMENT TEST LEVELS

In all three evaluation years, the data clearly indicated that sample students in Basic and Pilot Program schools had acute needs for compensatory education services, as evidenced by their scores on standardized achievement tests. For example, average scores on the reading and mathematics tests administered near the beginning of each year's school term ranged from the 13th to the 38th percentile. Viewing these results in terms of the test publishers' national norms, the typical ESAA student's performance fell below that of around three-fourths of the nation's students in the grade levels of interest. White sample students had higher average scores than Black or Spanish-background students in all three years, but were still below the national average. Females, and students of higher socioeconomic level, tended to have higher pretest scores regardless of race.

Although the sample groups' mean scores were well below the national average, there was a fairly wide spread in the scores of individual sample students. In Year Three, for example, around 8 percent of the students placed at the 85th percentile or higher in the national distribution. It should be noted in this connection that ESAA regulations do not include eligibility requirements at the individual student level. Thus, there is no explicit requirement that a student

have any specified degree of academic disadvantage in order to participate in ESAA activities.

Students in minority-isolated (Pilot) schools tended to have lower achievement scores than those in desegregated (Basic) schools, even allowing for differences in racial/ethnic composition of the two samples. This pattern of differences was quite consistent across both White and Spanish-background students, but was not found for Black students. It is not clear whether the overall achievement differences between Basic and Pilot schools are attributable in any causal fashion to differences in school environment, or whether they simply reflect the higher socioeconomic status of students enrolled in Basic schools.

IV. SCHOOL FUNDING PATTERNS

PROCEDURES FOR ESTIMATING PER-PUPIL EXPENDITURES

It was not possible, without placing an unreasonable respondent burden on the sample schools and districts, to collect data on the per-pupil expenditures for individual students in the evaluation sample. Furthermore, in the first two years of the evaluation, data were not available on the extent of overlap among the students participating in supplemental activities supported by different funding sources. It was not known, for instance, whether ESAA students in a given sample school were a subset of the students receiving ESEA Title I support, or whether different sets of students tended to receive support under different programs such as ESAA, Title I, state-supported programs, etc. In the absence of such information, per-student supplemental expenditures were calculated in Years One and Two by simply dividing each category of funds in a school by the total school enrollment. It was recognized at the time that this compromise procedure, though providing some consistency in approach across different expenditure categories, could not provide a realistic picture of actual per-pupil supplemental expenditures for the students represented by our evaluation samples (i.e., for ESAA-participating students in the treatment schools, and ESAA-eligible students in the control schools).

In Year Three (1975-76), new items were added to the district questionnaires, asking for the percentages of overlap among students receiving support from different supplemental funding sources. This information, combined with data on the funding level for, and the number of students participating in, each type of supplemental program, made it possible to compute reasonably accurate estimates of total per-pupil supplemental expenditures for ESAA-participating and ESAA-eligible students. Because the third-year calculations of per-pupil costs are so much more meaningful than those for the earlier years, only the Year Three expenditure data are reported in this overview.

RESULTS

Table 3, below, summarizes the expenditure patterns in sample ESAA-participating (treatment) schools. Each table entry represents the per-pupil expenditures (dollars and percentage of total) for the specified sample and expenditure category, averaged across all ESAA-participating students in the sample treatment schools. For these ESAA-participating students, ESAA provided by far the largest share of supplemental funding. ESAA expenditures accounted for roughly one-fourth of all per-pupil expenditures in the two elementary-level samples, but less than one-tenth of the total expenditures at the secondary level.

Table 3. Per-Pupil Funding in Sample ESAA-Participating Schools in 1975-76

Funding Source	Evaluation Sample		
	Basic Elementary	Basic Secondary	Pilot Elementary
ESAA	321 (21%)	91 (8%)	329 (25%)
Non-ESAA Supplemental	196 (12%)	3 (0%)	91 (7%)
Regular	1032 (67%)	1022 (92%)	916 (68%)
TOTAL	1549 (100%)	1116 (100%)	1336 (100%)

COMPARISON OF EXPENDITURES AT ELEMENTARY AND SECONDARY LEVELS

Total per-pupil expenditures were substantially lower at the secondary level than at the elementary level, because of the greater concentration of supplemental funds in elementary schools. This finding is somewhat surprising, as secondary-level per-pupil expenditures are normally expected to be higher within a given district than those at elementary schools in the same district. One explanation is that sample high schools were more likely than elementary schools to list their total enrollments as participating in a supplemental program; this automatically reduced the calculated per-student expenditure figures. Another possibility is that sample secondary schools might have been drawn from less affluent districts than sample elementary schools.

COMPARISON OF EXPENDITURES IN TREATMENT AND CONTROL SCHOOLS

An implicit assumption in the study's impact analysis design (i.e., in the comparison of outcomes for the paired treatment and control schools) was that, because of their ESAA participation, the treatment schools would have a significant advantage over control schools in total per-pupil expenditures. Further, it was assumed that this difference in total per-pupil expenditures would be roughly equal in size to the per-pupil ESAA expenditures in the treatment schools. Actually, in Year Three, these assumptions were met only for the Basic Elementary sample. Treatment schools in that sample had an average total per-pupil expenditure of \$1549, compared with only \$1169 in the paired control schools; this difference is both statistically significant and educationally meaningful. Treatment-control differences were much smaller and not statistically significant in the Basic Secondary sample (\$1116 for treatment schools versus \$1025 for control schools) and the Pilot Elementary sample (\$1336 versus \$1206).

In the case of the Basic Secondary sample, per-pupil ESAA expenditures were simply too small to have much influence on the total expenditures. In the Pilot Elementary sample, the ESAA expenditures in treatment schools were partially offset by the control schools' greater expenditures from non-ESAA supplemental sources (\$91 in treatment schools compared with \$130 in control schools). More importantly, there was considerable leakage of ESAA funds into the Pilot control group, because of concessions made to the sample Pilot districts in an effort to keep them in the longitudinal sample for the final evaluation year. The average per-pupil ESAA expenditure in Pilot "control schools" was over half as large as that in the paired treatment schools (\$168 compared with \$329). Theoretically, none of the control schools' ESAA funds were directed toward any students participating in the evaluation. In a few districts, control schools were allowed to spend those funds at the grade levels of interest to the study (i.e., 3, 4, or 5), but not for the sample students. Other districts applied ESAA dollars to different grade levels (i.e., 1, 2, or 6). In either case, however, it is unrealistic to assume that the sizable infusion of ESAA funds into Pilot control schools would have had no effect on the sample students. If

nothing else, the ESAA funds might have freed non-ESAA supplemental funds to benefit the sample students. It seems clear that in Year Three the intended funding distinction between treatment and control schools was weak for the Pilot Elementary sample, as well as for the Basic Secondary sample.

V. RESOURCE ALLOCATIONS AND SERVICES DELIVERED

ALLOCATIONS OF SUPPLEMENTAL FUNDS TO DIFFERENT ACTIVITIES

Percentage allocations and actual dollar allocations of supplemental funds expended for different types of program services and activities were determined for each year's samples. Again, the present discussion will focus on the Year Three Data, because of the improved procedures used in that year to estimate per-pupil expenditures.

The percentage allocations for the Year Three treatment (ESAA-participating) schools are shown below in Table 4. Overall, around two-thirds of all supplemental funds were allocated to basic skills instruction; this category includes relevant teacher inservice training as well as actual instructional services. The highest concentration of expenditures for reading and mathematics instruction was in the Basic Elementary sample (77 percent), while the other two samples each allocated around 64 percent of their supplemental funds for that purpose. At the high school level, the second-best-funded activity was intergroup relations and cultural enrichment (15 percent), while at the elementary level it was new curricula (5 percent and 19 percent for the Basic and Pilot samples, respectively).

For both elementary samples, ESAA funds were allocated in about the same proportions as non-ESAA supplemental funds among the different activities. At the secondary level, however, the allocation patterns for the two funding sources were substantially different in some areas. ESAA funds in the sample high schools were concentrated much more heavily than non-ESAA funds on intergroup relations and cultural enrichment activities (24 percent versus 0 percent), and less heavily on reading instruction (40 percent compared with 67 percent).

Table 4. Percentages of Supplemental Funds Allocated to Different Types of Services

Type of Funding	Reading	Math	Guidance and Counseling	Health Services	New Curricula	Intergroup and Cultural Enrichment	Career Education	Community Activities	Administrative and Auxiliary Services	Other Costs
ESAA	52.0	18.2	7.3	0.0	6.5	1.2	0.0	0.7	5.8	8.2
Non-ESAA	61.9	18.1	1.1	0.5	8.4	0.0	0.3	0.3	3.4	6.1
Total	56.3	21.0	2.6	0.2	5.3	0.4	0.1	0.8	4.7	8.6
ESAA	39.8	14.7	1.9	0.0	0.0	23.5	2.2	4.5	6.3	7.2
Non-ESAA	67.3	12.3	0.0	0.0	13.6	0.0	0.0	0.0	4.5	2.3
Total	48.8	15.7	1.7	0.0	0.5	14.5	1.8	4.5	5.5	7.0
ESAA	54.5	13.4	0.1	0.1	16.5	0.0	0.3	2.4	7.7	5.2
Non-ESAA	48.7	17.2	4.2	1.3	10.3	0.3	0.7	2.0	4.9	10.5
Total	51.6	12.7	1.6	0.4	19.2	0.1	0.3	1.9	6.2	5.9

BASIC
ELEMENTARY
SAMPLE

BASIC
SECONDARY
SAMPLE

PILOT
ELEMENTARY
SAMPLE

COMPARISON OF ALLOCATIONS IN TREATMENT AND CONTROL SCHOOLS

The previous section of this overview noted that, in the Basic Elementary sample only, treatment schools had significantly larger total per-pupil expenditures than control schools. Further analyses were performed to determine whether the treatment schools in the sample took advantage of their higher overall funding levels to focus greater resources and services on activities related to this study's major outcome measures. Such a trend was found in the Basic Elementary sample, where treatment schools averaged \$398 per pupil for reading and mathematics instruction, compared with only \$240 in the paired control schools; this difference is statistically significant. In the Basic Secondary and Pilot Elementary samples, treatment-control differences in allocations for basic skills instruction were much smaller, and failed to reach statistical significance.

Similarly, only the Basic Elementary sample showed any clear, consistent patterns of treatment-control differences in staff resource utilization for basic skills instruction. In particular, the treatment schools in that sample made significantly greater use of remedial specialists and instructional aides for regular reading instruction. Furthermore, average pupil/teacher ratios were significantly lower in treatment schools than in control schools, which suggests greater overall concentration of staff time in the treatment school classrooms.

TARGETING OF RESOURCES TO MEET NEEDS

There is considerable evidence that program resources and services in ESAA-participating districts, were, on the whole, effectively targeted toward schools and students with needs for those resources and services. It was clear in all three evaluation years that ESAA grants were awarded to districts with large percentages of disadvantaged and low-achieving students; those districts, in turn, targeted ESAA expenditures and services toward schools with large percentages of such needy students. In Year Two, higher levels of supplemental expenditures were found in schools with greater needs (i.e., lower achievement pre-test scores). Furthermore, evidence from both Year Two and Year Three indicated that greater program resources were focused on individual students with greater

academic needs. Specifically, in Year Two, students with lower achievement pretest scores tended to receive larger periods of exposure to basic skills instruction. In the Year Three Basic Elementary sample, more per-student hours of staff time were allocated to the needier students. For both the Basic Elementary and Secondary samples, moreover, the treatment schools showed better targeting of staff resources to needy (low-achieving) students than did the paired control schools, indicating a possible beneficial affect of the treatment school's participation in the ESAA program.

VI. ESAA PROGRAM IMPACT ON STUDENT ACHIEVEMENT

CROSS-SECTIONAL (ONE-YEAR) RESULTS

In the first evaluation year (1973-74), there was little evidence of substantial differences in program resources between the ESAA-funded (treatment) schools and the non-ESAA-funded (control) schools. For this reason, and because the sample students had received only about five and a half months' exposure to ESAA programs that were still in their initial implementation phase, it came as no surprise that there was no evidence of ESAA program impact, i.e., that no treatment-control differences in student achievement test gains were found.

During the second year (1974-75), there were indications of a possible trend toward larger differences in total per-student funding between ESAA-participating and non-ESAA-participating schools, though allocations for activities seemingly relevant to the outcome measures (e.g., reading and math instruction) were still similar for the two groups. Paralleling the growth in size of the treatment-control funding differential, there appeared to be stronger evidence of ESAA program impact, at least for one of the samples. Within the Basic Elementary sample, treatment (ESAA-funded) schools made larger gains than control (non-ESAA-funded) schools in both reading and mathematics, and in all three grade levels sampled. Despite the consistency in direction of these findings, however, none of the treatment-control differences in the Basic Elementary sample was statistically significant. The summary report for Year Two* concluded that, "it will be important in subsequent analyses to examine results for the third evaluation year (1975-76) to determine whether the apparent trends toward greater treatment-control differences in resources and in outcomes increased in the final year of the evaluation."

*National Evaluation of the Emergency School Aid Act (ESAA): Summary of the Second-Year Studies. System Development Corporation, TM-5236/011/00, July, 1976.

In the Year Three Basic Elementary sample, as we have noted above, there was a statistically significant and educationally meaningful advantage for treatment schools over control schools in total per-pupil funding, in per-pupil supplemental allocations for reading and math instruction, and in the use of remedial specialists and instructional aides to augment staff resources in regular reading classes. By contrast, there were no consistent patterns of significant treatment-control differences in resources or services for the Basic Secondary sample or the Pilot sample. From these findings, the following conclusions might reasonably be drawn:

- The Year Three Basic Elementary sample met the assumption implicit in the study's design, i.e., that there would be substantial treatment-control differences in funding and services. Since those differences showed evidence of increase from one evaluation year to the next, it might be expected that evidence of ESAA program impact (treatment-control differences in achievement gains) should be stronger in Year Three than in previous years.
- The Year Three Basic Secondary and Pilot Elementary samples did not meet the impact design's assumptions regarding treatment-control resource differences. Thus, no significant difference in achievement gains for treatment and control schools should be expected.

In general, the Year Three findings confirmed these expectations. Table 5, below, shows the pretest scores, posttest scores, and gains in reading and mathematics for each grade of the Year Three Basic Elementary sample.*

At the third grade, the control-group students had significantly higher pretest scores than those in the treatment schools. Where such initial inequities exist, comparisons of gains are generally not meaningful or interpretable. In grades

*All scores in the table are expressed in terms of special scale scores. The derivation of these scores is described in the ESAA Year Three report, referenced at the end of this document.

Table 5. Year Three Analysis of Achievement Outcomes: * Basic Elementary Sample

Grade Level	Number of Pairs	Group	Reading			Mathematics		
			Pretest	Posttest	Gain	Pretest	Posttest	Gain
3	21	Treatment	488.7	527.2	38.5	443.0	504.2	61.2
		Control	513.1	551.4	38.3	459.3	523.9	64.6
		T-C Difference	-24.4**	-24.2**	0.2	-16.3**	-19.7	-3.4
4	22	Treatment	543.3	579.0	35.7	526.3	573.2	46.9
		Control	562.4	590.2	27.8	541.8	578.8	37.0
		T-C Difference	-19.1	-11.2	7.9	-15.5	-5.6	9.9**
5	21	Treatment	597.3	623.9	26.6	598.1	634.9	36.8
		Control	617.3	635.5	18.2	611.2	642.0	30.8
		T-C Difference	-20.0	-11.6	8.4**	-13.1	-7.1	6.0

*Achievement outcomes are expressed in units of the ESAA scale scores.

**A statistically significant treatment-control difference ($p \leq 0.05$ level) is present. Positive differences indicate higher treatment scores than control scores, and negative differences indicate higher control scores.

4 and 5, however, there was a clear pattern of larger gains for treatment schools than for control schools. Two of the four differences (fourth-grade mathematics and fifth-grade reading) were statistically significant ($p \leq .05$);* that is, differences as large as these would be expected to occur by chance fewer than five times in a hundred. Furthermore, a third difference (fourth-grade reading) almost reached significance ($p = .06$). These treatment-control outcome differences at grades 4 and 5 appear educationally meaningful, as well; they are roughly equal to two or three percentile points, relative to the test publisher's national norms. It should be further noted that the treatment-control outcome comparisons were performed at the school level, and involved a maximum of only 22 pairs of schools; thus the statistical significance of the differences for fourth-grade math and fifth-grade reading is not simply an artifact of large sample sizes in the analyses.

POSSIBLE EXPLANATIONS FOR YEAR THREE IMPACT FINDINGS

Additional analyses were performed to determine whether the stronger evidence of program impact in the Basic Elementary sample for Year Three might have resulted in part from selective project attrition. It was felt that weaker, less successful ESAA projects might have been eliminated from further funding at the end of Years One and Two, either by deliberate screening on the part of USOE or by some self-selection process. However, statistical comparisons of projects continuing in the sample, versus those dropping out at the end of Years One and Two, failed to show evidence of such selective attrition.

*It will be noted that the treatment schools' pretest scores for grades 4 and 5 were consistently lower than those of the control schools, though the differences were not statistically significant. To ensure that the evidence of ESAA program impact for those grades was not simply an artifact of initial treatment-control differences, the impact analyses were repeated using posttest scores adjusted for pretest level, race, and socioeconomic indices. In addition, separate sets of impact analyses were performed for schools with different levels of pretest scores. Neither of these reanalyses altered the basic impact findings reported here.

Two other explanations for the increased treatment effect in Year Three seem more probable. First, as noted earlier, the Basic Elementary treatment schools' funding advantage over the control schools was greater in Year Three than in previous years; this could have strengthened program impact. Second, the Year Three Basic Elementary treatment schools may have benefited from program maturation. New programs, such as ESAA, necessarily devote a large part of their effort in early years to planning and reorganization, staff recruitment and training, and development of essential materials. In later years of a project's life, these start-up activities can be reduced greatly in favor of activities more immediately and directly beneficial to student achievement, such as actual classroom instruction, and preparation and implementation of daily lesson plans.

The program maturation hypothesis could not be directly tested in this study. However, if such maturation does account in part for the stronger evidence of program impact in Year Three, this has an important implication for future program evaluations. It would add considerable weight to the arguments of many program administrators, that the merits of their programs should not be judged on the basis of outcome data collected in the first year or two of those programs' existence.

As expected, given the lack of significant funding and resource differences between treatment and control schools in the Basic Secondary and Pilot Elementary samples, those two samples showed no evidence of program impact with respect to basic achievement.

LONGITUDINAL (THREE-YEAR) RESULTS

Analyses were also performed to compare achievement growth for the treatment and control longitudinal samples, i.e., for the cohorts of students who were in the sample for all three years of the evaluation. In general, there was no clear evidence of cumulative ESAA program impact over the three years of the evaluation for any of the longitudinal samples, as the treatment and control groups made similar achievement gains. For most combinations of sample and

achievement outcome measures, the control groups started slightly higher (though in general not significantly so), and stayed slightly higher over successive test administrations, with little sign of any "catch-up" by the treatment group. One possible exception was in reading achievement for the Basic Elementary sample. Here, the gap between treatment and control schools was reduced between pretesting and posttesting in Year Three. Although this convergence trend was not statistically significant, it seems consistent with the finding of ESAA program impact for the Year Three Basic Elementary sample.

VII. RELATIONSHIPS AMONG PROGRAM VARIABLES, STUDENT ACHIEVEMENT, AND SCHOOL CLIMATE

ANALYTIC METHODOLOGY

For the exploratory analyses reported here, data from ESAA-funded (treatment) and non-ESAA (control) schools were combined, as the purpose was to determine what program features were associated with successful outcomes, regardless of funding sources for those programs. The relationships discovered in these analyses appear to have important policy implications for ESAA and similar educational programs.

Some of the results discussed below are based upon the entire samples for the overall evaluation, while others are based on the in-depth study subsamples (24 elementary schools in Year Two, and 26 elementary schools in Year Three). Many of the in-depth analyses used program description data that were collected only in the in-depth sites, e.g., observations by SDC personnel of classroom instructional procedures and of resource utilization.

Relational analyses conducted prior to Year Three were hypothesis-seeking in nature; that is, fairly large numbers of program variables were investigated in an effort to identify the subset of variables most strongly related to student achievement. The danger in such an approach is that statistically significant but totally spurious relationships can occur by chance, simply because of the large number of analyses performed. For this reason, an effort was made in the Year Three analyses, for both the overall evaluation and the in-depth study, to focus primarily on variables that had been found significantly related to achievement in the preceding years. Thus, the Year Three analyses were more confirmatory in nature, and less subject to the dangers of spurious relationships. Even so, the relational analyses for all three years were quasi-experimental rather than being based on true experiments, as neither schools nor students were randomly assigned to the various program approaches being compared. A relationship attributed to some program variable may actually result from

sample bias or from uncontrolled variations in other environmental factors. For this reason, in the present discussion of results, we have chosen to emphasize patterns of relationships that appear to have some consistency across years. Somewhat greater confidence can be placed in such findings, particularly where they are in agreement with results of earlier evaluation studies.

The following paragraphs present the results of relational analyses, first for the overall evaluation, and then for the in-depth study. It should be noted that the trends discussed here were not necessarily found for every grade level or for every subsample. Readers interested in more detailed information about the findings are referred to the individual years' evaluation documents listed at the end of this summary report.

RESULTS FROM OVERALL EVALUATION

The categories of program variables most consistently related to achievement were (1) per-pupil level of program funding, (2) hours of instruction experienced by students, (3) pupil/teacher ratio, and (4) district or school activities supporting desegregation. Less consistent patterns were found for teacher inservice training and for certain specific instructional practices.

Funding and Allocation Levels

In Year One, significantly larger reading and mathematics gains occurred in Pilot Elementary schools with higher per-pupil levels of supplemental funding allocations for reading instruction. In Year Two, larger mathematics gains were found in Pilot Elementary schools with higher per-pupil supplemental funding allocations for mathematics instruction; furthermore, for both the Pilot Elementary and Basic Elementary samples, reading and mathematics achievement tended to be higher in schools with higher per-pupil levels of regular funding and of total (regular plus supplemental) funding. Similarly, in Year Three, for the two elementary samples, larger gains in reading and/or mathematics were associated with higher per-pupil levels of supplemental expenditures and of

total expenditures. Overall, it appears that expenditure and allocation levels, particularly with respect to supplemental funding sources, were influential in determining achievement outcomes at the elementary levels, but the evidence is equivocal at the secondary level.

Period of Exposure to Instruction

For all three years, at least one of the evaluation samples showed a positive and statistically significant relationship between the amount of time students were exposed to instruction and their achievement gains. In Year One, the Pilot Elementary sample showed larger reading gains where students received more hours of reading instruction. For the same sample in Year Two, and for both the Basic Elementary and Pilot Elementary samples in Year Three, a similar relationship was shown between mathematics exposure and gains in mathematics achievement scores.

Pupil/Teacher Ratio

In both Years One and Two, there was some evidence of significantly larger mathematics achievement gains in schools with smaller pupil/teacher ratios. Such a relationship was found in Year One for the Pilot Elementary sample, and in Year Two for the Basic Secondary sample.

Program Activities Supporting Desegregation

Statistically significant relationships were found in each of the evaluation years between student achievement and district or school activities designed to facilitate and support desegregation (e.g., media presentations, public meetings, and staff meetings concerned with interracial interactions). In the Year One Basic Elementary sample, larger achievement gains occurred in schools with greater frequency of such activities, and similar relationships were found for the Basic Secondary sample in both Years Two and Three. Surprisingly, the Year Three Basic Elementary sample showed a negative relationship between achievement gains and level of desegregation-supporting activities. One possible

explanation for this apparent contradiction related to the amount of time the schools had been desegregated. In Year One, desegregation had been fairly recent in many Basic Elementary schools; in that same year, a positive relationship was found between desegregation-related activities and achievement. In Year Three, when most of the sample schools had been desegregated for several years, the relationship became negative. It seems quite possible that desegregation-related activities in Year One represented positive district and school attitudes toward desegregation, whereas by Year Three, only schools with severe integration problems--problems that could adversely affect achievement--were still devoting an appreciable effort to the direct support of desegregation. Secondary schools, by contrast, may have desegregated more slowly; in such a case, desegregation-supporting school activities in Year Three may still have represented a positive and affirmative action rather than a reaction to obvious disharmony.

Teacher Training

Little relationship was found between the teachers' preservice training and student achievement, but in Year Two there was some evidence that inservice training might be influential. Specifically, for the Basic Elementary sample, larger reading gains were found in schools whose teachers had more recent inservice training. In the Basic Secondary sample, students of mathematics teachers with more specialized training in mathematics instruction tended to make larger gains in mathematics achievement.

Instructional Practices

Very few statistically significant relationships were found between specific classroom teaching practices and student achievement. However, in the Year One Basic Elementary sample, larger mathematics achievement gains were associated with more frequent use by teachers of diagnostic tests. Also, in the Year Two Basic Secondary sample, larger mathematics gains were found where teachers placed greater emphasis on drill and practice procedures. A possible common factor in

these findings is that, in both cases, a more structured, less intuitive approach to classroom instruction seems to have produced better achievement gains.

RESULTS FROM IN-DEPTH STUDY

Analyses based on the in-depth study samples of elementary schools yielded additional findings of considerable importance for ESAA program policy formulation. In these analyses, elementary schools found successful in reading or math were compared with other, less successful schools. The major dimensions on which comparisons were based were the schools' organizational climate, parental and community involvement, reading and mathematics instructional practices and related teacher attitudes, reading and mathematics instructional resources, and equality of educational opportunity. In-depth study results in each of these areas are summarized below.

Organizational Climate

In both years of the in-depth study (1974-75 and 1975-76), the schools' organizational climate was a significant factor in their ability to improve students' achievement. The most consistent finding, in both years, was that schools characterized by strong administrative leadership were significantly more likely to raise students' achievement levels. For purposes of these analyses, a school was rated high in administrative leadership if it had the following characteristics:

- The principal felt strongly about instruction, had definite ideas about proper instructional practices, and promoted a point of view.
- The principal effectively communicated his/her viewpoints about instruction, as evidenced by the teachers' ability to report those viewpoints accurately. Specific communication techniques found successful were regular reviews and discussion of teachers' instructional performance by principals; faculty meetings where instructional methods were discussed; and in-school meetings on educational theories and concepts.

- School administrators, especially the principal, took the major responsibility for decisions concerning the selection of the school's basic instructional materials, the planning of school programs, and the evaluation of programs. In the less successful schools, such decisions were more often left to the individual teachers.
- Administrators emphasized academic standards by reviewing teaching performance, by opposing the postponement of basic skills instruction, and by requiring students to repeat grades when standards for achievement were not met.

Reading and Mathematics Instructional Practices

In the first year of the in-depth study, a higher probability of school success in improving achievement was found in schools whose teachers:

- Organized their instruction around lesson objectives defined in terms of specific skills and knowledge to be mastered by the students,
- Gave students practice sessions incorporating many of the steps necessary for mastery of the lesson objectives, and
- Praised the students less frequently.

Although it may seem surprising that greater praise was associated with lower achievement, there were indications that the teachers who made more lavish use of praise were somewhat non-selective in that use--that is, they tended to give praise regardless of whether the students were demonstrating learning behaviors. In addition, such teachers tended to stress psycho-emotional goals as more important than academic goals, and also tended to give the students less practice time. Thus, extensive use of praise may have been associated with a less task-oriented approach to instruction.

Efforts in the second in-depth study year to confirm findings of the previous year were generally successful. That is, most variables related to achievement in the first year were also important in the second year. In the second year,

however, there was a tendency for combinations of factors rather than individual variables to show statistically significant associations with school success. Specifically, a school's students were significantly more likely to make gains in math achievement under the following conditions:

- When teachers praised and rewarded students frequently and felt it important to restrict awards to occasions when students made academic progress. Although this finding may appear to contradict the earlier year's negative association between praise and achievement, it seems on closer inspection to augment the previous result by giving insight into the conditions under which praise can have positive or negative effects. A reasonable, if oversimplified, summary of the relationship is that the use of carefully structured praise on a contingent basis to reward desired behaviors will tend to improve achievement, while liberal but non-selective and unstructured use of praise is more likely to have a negative effect on achievement.
- When teachers used behavioral objectives and attached importance to setting challenging goals. Again, this finding appears to amplify previous in-depth study findings by defining conditions under which the teachers' use of behavioral goals may benefit achievement growth. Teachers increase the probability of student achievement gains by setting such behaviorally defined objectives, but only if the incremental step represented by each new objective places intellectual demands on the students.
- When teachers reported that they placed strong emphasis on individualized instruction and on the use of behavioral objectives. This finding suggests that individualized instruction may be especially effective when teachers establish explicit behavioral goals.

A central theme underlying all of these findings is that students had a better chance of academic growth when the teachers strongly structured many aspects of the classroom environment, including their use of social reinforcers as well as

their direct instructional procedures. Other in-depth analyses strongly suggest that the mediating factor in this relationship was the degree of student task-orientation, i.e., the degree to which students were observed to be attending to their studies. It appears that students in more strongly structured learning situations were more task-oriented, and that more task-oriented students tended to have larger achievement gains.

Resource Utilization

In the first year's in-depth study, schools were significantly more likely to be successful in raising mathematics achievement when:

- More of their resources were allocated to remedial specialists, and
- Less of their resources were devoted to non-certificated instructional aides.

The second year's in-depth data failed to confirm these specific relationships, but did yield other findings related to staff development and utilization. Greater probability of success in mathematics was found in schools where:

- Principals had acquired more inservice training through participation in staff development activities.
- Professional development activities for the staff, such as weekend courses, inservice training, help in covering costs of attendance at professional conferences, etc., were more actively encouraged.
- Remedial specialists were used in a greater variety of ways, e.g., referrals, on-call basis, conferences with teachers, etc.

Equality of Educational Opportunity

Equal treatment for all students, regardless of their race or ethnicity, is both a moral and a legal requirement; it requires no further justification on the basis of possible effects on student achievement. It is of interest, nevertheless, that both years' in-depth studies yielded some evidence of positive

relationships between equality of educational opportunity (EEO) and achievement. In the first year, reading achievement gains were significantly less likely to be found in schools whose teachers were observed directing a disproportionate amount of negative feedback toward minority-group students (e.g., pointedly ignoring the students, criticizing them, or sending them out of the classroom). The second year's findings showed greater likelihood of reading and mathematics achievement gains in schools that placed greater emphasis on cultural enrichment activities (as evidenced, for example, by the principal's emphasis on intergroup relations activities, by the use of multi-ethnic reading materials and school displays, and by the amount of library material on cultural enrichment/inter-group relations themes).

Also found in both years was a strong agreement among the different indicators of the schools' degree of commitment to EEO. If an in-depth school used fewer multi-ethnic instructional materials, it was also more likely to have segregated seating patterns, and its teachers were more likely to criticize minority students disproportionately. Conversely, a school that was strong in one dimension was likely to be strong on all indicators.

Parent and Community Involvement

Data from the first year of the in-depth study showed greater likelihood of improving students' mathematics performance in schools where parents were present in the classrooms as paid instructional aides, volunteers, or visitors. No such relationship appeared in the second year, however.

VIII. SUMMARY OF FINDINGS, AND RELATIONSHIPS TO PREVIOUS STUDIES

The data from all three years of the ESAA evaluation show clearly that dollars were targeted toward needy school districts, schools, and students; further, those dollars were translated into services directed toward students with needs for such services. This finding, encouraging in its own right, leads logically to the question of program impact: given the needs of the ESAA students, as evidenced by their low achievement levels and low socioeconomic status, and given the apparently effective targeting of resources within ESAA-funded schools, did ESAA participation help to increase student achievement?

IMPACT RESULTS

For elementary schools participating in the ESAA Basic Program, the evidence of positive ESAA program impact was progressively stronger in each new year of the study. In Year One (1973-74), there was virtually no indication that ESAA participation might be having a beneficial effect on achievement. This was not a surprising finding, considering the brief period of student exposure (under six calendar months) to an ESAA program that was still in its early stages of implementation, and the fact that differences in total per-pupil funding for the ESAA-participating (treatment) and non-ESAA-participating (control) schools were fairly small. In Year Two (1974-75), the treatment-control differences in total per-pupil expenditures increased slightly, and many of the ESAA-participating schools (those funded for Year One as well as Year Two) had had more time to get their programs organized. Although there was still no statistically significant evidence of ESAA program impact for Year Two, there was an evident trend; ESAA-funded schools made larger achievement gains than non-ESAA-funded schools in both reading and math, for all three elementary grade levels tested (grades 3, 4, and 5). Finally, the Year Three data showed clear-cut evidence of program impact, as there were statistically significant differences in achievement gains, favoring the ESAA-funded schools, for both reading and mathematics. This increase in program impact, relative to previous years, probably reflects the growing maturation of the ESAA projects, and the fact that treat-

ment-control differences in resources were greater for Year Three than in previous years; in Year three, specifically, ESAA-funded schools had statistically significant advantages over non-ESAA-funded schools in total per-pupil funding, in allocations of supplemental funds for reading and mathematics instruction, and in remedial staff time devoted to such instruction.

No statistically significant evidence of ESAA program impact was found in any of the three years for secondary schools participating in the Basic Program, or for elementary schools in the Pilot Program. It should be noted, however, that in most instances, dollar allocations for reading and mathematics instruction in ESAA-funded (treatment) schools did not differ significantly from those of the control schools in the Pilot Elementary and Basic Secondary samples. Neither was there any clear, consistent evidence of different levels or types of staff utilization for basic skills instruction in ESAA-funded and non-ESAA-funded schools. Without such treatment differences, it would seem unreasonable to expect significant outcome (achievement) differences.

RELATIVE EFFECTIVENESS OF DIFFERENT PROGRAM APPROACHES

While a major focus of the evaluation was on the overall impact of the ESAA programs, other important analyses attempted to assess the relative effectiveness of different program approaches or components, regardless of whether those components were paid for by ESAA or from other funding sources; for these analyses, data from the ESAA-participating and non-ESAA-participating schools were pooled. In some instances, fairly specific operational variables were consistently related to student achievement. In certain other cases, the specific variables associated with achievement might change from year to year, but some of those variables formed patterns involving more general, underlying dimensions of the program. Based on a combination of findings from the overall evaluation and the in-depth study, the more pervasive trends with potential policy implications for future program decisions may be summarized as follows:

- Strength of administrative leadership was a major factor in the schools' ability to improve student achievement. There was consistent evidence that improved achievement was more likely to be found in schools whose principals (a) felt strongly about instruction; (b) effectively communicated their viewpoints about instruction to teachers, through principal/teacher discussions, reviews of teaching performance, etc.; (c) took a dominant role in decisions about the selection of instructional materials and in program planning and evaluation; and (d) emphasized academic standards, by opposing the postponement of basic skills instruction and by requiring students who did not meet academic standards to repeat grades.

These findings are consistent with those of several other evaluation studies in stressing the importance of the principal's role. For example, a recent in-depth study of compensatory reading programs (Trisman, Waller, and Wilder, 1975) reported that one of the critical characteristics of the unusually effective programs being studied was strong administrative leadership. Another study, investigating reading gains in Los Angeles schools with heavy concentrations of minority-group students, concluded that, "Among school-level factors that affect reading achievement, the leadership role of the principal is one of the most important" (Armor, Conry-Osequera, Cox, King, McDonnell, Pascal, Pauly, and Zellman, 1976). There is also some corroborative evidence, of an informal nature, for the importance of emphasizing academic standards. For example, a national weekly, The Observer, recently described dramatic improvements in student achievement for Greenville County (Virginia) schools that replaced "social promotions" with "competency promotions" (Greene, November 6, 1976). In two years after the policy change on promotions, that county's sixth-grade students rose from last place among Virginia schools to an average or above-average position in academic achievement.

- In general, achievement gains tended to be larger in schools with greater resource availability and utilization. There were some variations over years and across samples in the specific resource indices that were significantly related to achievement. For example, the first year's in-depth study showed a positive association between student achievement and the school's level of use of remedial specialists, while in the next year's study, achievement was related to the number of different ways in which those specialists were utilized (e.g., referrals by teachers, regular periodic assignments of students to specialists, assisting in regular classroom instruction as needed). Similarly, in the overall evaluation, analyses of data for different evaluation years and different samples have shown achievement to be related in some instances to per-pupil levels of supplemental expenditures for reading or math instruction, and in other cases to the regular or total (regular plus supplemental) levels of per-pupil expenditures. Overall, however, expenditures and resource allocations were evidently influential in determining achievement at the elementary levels, particularly when those expenditures and resources were focused on activities directly relevant to reading and math instruction.

Also pertinent to the issue of resource utilization was the finding, in two of the study years, that achievement gains were greater in schools with smaller pupil/teacher ratios; this can be interpreted as further evidence of the beneficial influence of concentrating greater staff resources, per student, in the classroom.

Previous studies have not agreed among themselves with respect to the importance of program funding and resources in determining student outcomes. For example, a recent cost-effectiveness study of a sample of Michigan elementary schools (the 1974-75 Michigan Cost-Effectiveness Study, 1976) found that the more successful sites had significantly higher total program costs and significantly greater per-pupil costs

for classroom reading activities. On the other hand, another recent study, evaluating compensatory reading programs, found no relationships between per-pupil levels of resource costs for reading programs and achievement gains (Flynn, Hass, Al-Salam, 1976). Some of these differences in findings undoubtedly reflect variations in the study samples, methodology, and instruments, as well as in the measures of resource allocations employed. Clearly, it is not dollars as such that determine program success, but rather the way in which those dollars are administered and converted into enhanced resources and/or staff skills influencing the day-to-day classroom experiences of the students. It seems logical to assume that, where no relationship was found between program resources and student achievement, it was at least in part because the schools with greater resources failed to apply them in a manner that would usefully alter those classroom experiences.

- Efforts to promote equality of educational opportunity (EEO) were associated with greater probability of school success in raising student achievement levels. As with the other variables examined in these comparisons of different program approaches, the evidence here is somewhat circumstantial--i.e., the present study does not permit firm conclusions about the direction of causal relationships, if any, between EEO efforts and achievement. Nevertheless, results of the two years' in-depth studies strongly suggest that district or school activities aimed at cultural enrichment and at the improvement of intergroup relations can have beneficial effects on achievement.

There have been numerous studies in the past of relationships between desegregation and achievement, but most of those studies have looked only at gross statistical indicators of the degree of racial balance among schools in a district, rather than at specific actions being taken within schools to reduce segregation or to improve racial relationships. For this reason, perhaps, Nancy St. John's review of such studies (1975)

reported little agreement among them, and concluded that how desegregation was implemented was probably at least as important as the fact of desegregation itself. A more relevant study, in terms of the present ESAA evaluation, was one by Forehand, Ragosta, and Rock (1976). Several of the indicators used in that study to describe "teaching for effective race relations" were similar to those employed in the ESAA evaluation, e.g., use of multiracial curriculum materials, teaching of minority-group history and culture, etc. Forehand and his associates found that "teaching methods oriented toward good race relations are associated with improved achievement" (p. v).

- Students with more hours of exposure to reading or mathematics instruction tended to make significantly larger achievement gains in the content area concerned. This finding, though hardly surprising, has important policy implications, especially in light of the increasing pressures on many schools to substitute popular non-academic activities for some of the student time formerly devoted to basic skills instruction. An obvious conclusion is that, if a school or district places high priority on basic achievement, and if such achievement is to be a major criterion for judging program success, then that district or school should concentrate the bulk of its efforts and resources on basic skills instruction, and should attempt to maximize the students' exposure to such instruction.

Several previous studies have similarly yielded positive relationships between exposure time and achievement (Kidder, 1974; O'Reilly, 1974; Wiley and Harnischfeger, 1974). However, a recent study of a reading program in several Los Angeles schools failed to show beneficial efforts of increased exposure (Armor, Conry-Osequera, Cox, King, McDonnell, Pascal, Pauly, and Zellman, 1976).

- Schools where the instruction was more highly structured, and where greater demands were placed on the students' academic performance, had a better chance of improving student achievement. The in-depth study findings suggest that these dimensions of classroom practice were important determiners of achievement gains, and that they tended to interact in their influence. Though the specific relationships varied somewhat over the two in-depth study years, it seems reasonably clear that beneficial effects were derived from teachers' efforts to structure the students' learning experiences by using lesson objectives defined in terms of desired student behaviors; by giving frequent diagnostic tests to assess progress toward each learning goal; and by making extensive use of drill and practice procedures to help students achieve the defined objectives. Additional findings, particularly in the second year of the in-depth study, indicate that structuring of lessons may not have a positive influence on achievement unless sufficiently stringent demands are placed by the school on the students' level of academic effort and attainment. In the second year, for instance, teacher use of behavioral lesson objectives was beneficial provided each objective was sufficiently difficult to be challenging to the students. Similarly, there was some evidence that the liberal use of praise could have negative impact on achievement unless that praise was contingent on evidence of student progress toward learning goals. At the administrative level, moreover, greater chance of success was found where principals required low-achieving students to repeat grades.

In one sense, all of these findings can be viewed as examples of the importance of structure. By this interpretation, the concept of "structure" is applied to the larger school context; that is, it includes not only the planning and presentation of instructional content, but also the control of motivational factors by careful selection of the difficulty level of successive lesson objectives, and by arranging specific contingencies for praising students, for advancing them in grades, etc.

It should be remembered that the sample schools and students in this evaluation are not representative of the total range of American students, but tend to be at the lower end of the national distribution with respect to their socioeconomic status and achievement level; thus, results of the present study cannot necessarily be generalized to the total school population. Our findings are quite consistent, however, with those of several other studies involving disadvantaged students. A recent review, summarizing and synthesizing the results of a number of earlier studies, concluded that, "The tightly structured program approach including frequent and immediate feedback to the student combined with a tutorial relationship, individual pacing, and somewhat individualized programming are positively associated with accelerated pupil achievement" (Gordon, 1971, p. 24). Other pertinent evidence comes from a two-year study of Michigan schools (Michigan, 1976), which found degree of program organization to be positively associated with level of reading achievement. Also relevant are studies by Furst and Amidon (1967), Penny (1969), and Soar (1966), showing student achievement to be higher in classrooms where teachers presented overviews or cognitive frameworks to introduce lesson materials and to help the students place the new materials in a meaningful and structured context. Finally, in relation to the issue of academic standards and selective use of praise, a study of teacher standards in San Francisco high schools (Fernandez, Espinosa, and Dornbusch, 1975, p. 2), found lower achievement where the schools' evaluation system "substituted praise and warmth for challenging academic standards."

IX. THE EVALUATION DESIGN IN HINDSIGHT

In concluding this summary report, it may be useful to comment briefly on two or three basic issues concerning the design of the ESAA evaluation. First, a decision was made early in the study to place primary emphasis on the impact evaluation. This decision led to a sampling procedure based on the selection of pairs of similar schools within districts, and the random assignment of schools in each pair to the treatment and control conditions. The purpose of these procedures was to approach a true experimental design involving comparisons of outcomes in the treatment and control groups, so that confidence could be placed in the impact findings.

This design decision had both positive and negative consequences for the study's implementation and results. There were practical problems in getting the co-operation of certain school districts in a sampling plan that intentionally excluded some otherwise eligible schools (those assigned to the control condition) from receiving ESAA funding. These problems were not serious in the first evaluation year, but by Year Three, efforts to retain schools for the Pilot and Basic Secondary samples, which were small to begin with, led to compromise agreements with some districts that tended to degrade the funding distinction between treatment and control schools. Furthermore, the requirement that sample districts have at least one pair of similar schools designated for ESAA participation, tended to restrict the number of eligible districts for the samples and systematically eliminated very small districts from possible representation.

Despite these problems, however, the decision to maximize the rigor of the impact design by random assignment procedures seems in retrospect to have been a sound one. No other approach would have permitted as much confidence to be placed in the impact findings. Furthermore, it seems quite likely that a less rigorous design, such as the non-random assignment of "similar" schools from non-ESAA-funded districts to serve as the control group, would have obscured

the evidence of ESAA program impact found for the third-year Basic Elementary sample by confounding district-level and school-level differences between the treatment and control groups and by allowing other uncontrolled sources of sample variations. All in all, the special efforts made in this study to form comparable treatment and control groups seem well justified.

Hindsight also confirms the value of the in-depth examination of selected schools. The in-depth study effectively complemented the overall evaluation design, and provided detailed observational data about program operations that would not otherwise have been available. Thus, while the overall evaluation yielded important descriptive and impact data on the national samples of ESAA-participating districts and schools, the in-depth study was the major source of information about specific administrative and classroom factors contributing to success in local projects.

Another issue of some importance concerns the fact that statistically significant evidence for ESAA program impact in the Basic Elementary sample did not appear until the third year of program operations. The data from this and other studies appear to show that large new programs such as ESAA require at least two or three years to "shake down," and that evaluations prior to that time may do an injustice to the programs. This is not to suggest that no program evaluations should be performed in the first year or two after a program's creation. On the contrary, it can be valuable to begin immediately collecting baseline performance data (for comparison with later achievement) and information about program processes and implementation. Such formative evaluations can greatly benefit program development and improvement. It may be wise, however, to defer summative (impact) evaluation until at least the second or third year of program operations, if one wishes a true picture of the program's potential contribution.

EARLIER ESAA EVALUATION REPORTS

Combined ESAA Pilot and Basic Program Evaluations: Final Study Design. System Development Corporation, TM-5236/002/00, January 1974.

Combined ESAA Pilot and Basic Program Evaluations: Final Study Design, Volume II. System Development Corporation, TM-5236/007/00, August 1975.

Combined ESAA Pilot and Basic Program Evaluations: Final Study Design, Volume III. System Development Corporation, TM-5236/015/00, February 1977.

Ozenne, D. G.; Van Gelder, N. C.; and Cohen, A. J. Emergency School Aid Act (ESAA) National Evaluation: Achievement Test Restandardization. System Development Corporation, TM-5236/006/00, November 1974.

Coulson, J. E.; Ozenne, D. G.; Van Gelder, N. C.; Inuzuka, D.; Bradford, C.; and Doherty, W. J. The First Year of Emergency School Aid Act (ESAA) Implementation. System Development Corporation, TM-5236/008/00, September 1975.

Coulson, J. E.; Ozenne, D. G.; Bradford, C.; Doherty, W. J.; Duck, G. A.; Hemenway, J. A.; and Van Gelder, N. C. The Second Year of Emergency School Aid Act (ESAA) Implementation. System Development Corporation, TM-5236/009/00, July 1976.

Coulson, J. E.; Ozenne, D. G.; Hanes, S. D.; Bradford, C.; Doherty, W. J.; Duck, G. A.; and Hemenway, J. A. The Third Year of Emergency School Aid Act (ESAA) Implementation. System Development Corporation, TM-5236/014/00, March 1977.

Wellisch, J. B.; Marcus, A. C.; MacQueen, A. H.; and Duck, G. A. An In-Depth Study of Emergency School Aid Act (ESAA) Schools: 1974-1975. System Development Corporation, TM-5236/010/00, July 1976.

Wellisch, J. B.; Carriere, R. A.; MacQueen, A. H.; and Duck, G. A. An In-Depth Study of Emergency School Aid Act (ESAA) Schools: 1975-1976. System Development Corporation, TM-5236/013/00, March 1977.

Coulson, J. E. National Evaluation of the Emergency School Aid Act (ESAA): Summary of the Second-Year Studies. System Development Corporation, TM-5236/011/00.

REFERENCES

Armor, D.; Conry-Osequera, P.; Cox, M.; King, N.; McDonnell, L.; Pascal, A.; Pauly, E.; and Zellman, G. Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools. Santa Monica, California: Rand Corporation, August 1976.

Fernandez, C.; Espinosa, R.; and Dornbusch, S. M. Factors Perpetuating the Low Academic Status of Chicano High School Students. Palo Alto, California: Stanford Center for Research and Development on Teaching, 1975.

Flynn, D.; Hoss, A.; and Al-Salam, N. An Evaluation of the Cost Effectiveness of Alternative Compensatory Reading Programs (Volume III). Bethesda, Maryland: RMC Research Corporation, April 1976.

Forehand, G. A.; Ragosta, M.; and Rock, D. A. Conditions and Processes of Effective School Desegregation. Princeton, New Jersey: Educational Testing Service, 1976.

Furst, N. F., and Amidon, E. "Teacher-Pupil Interaction Patterns in Elementary School," in E. Amidon and H. Hough (eds.), Interaction Analysis Theory, Research and Application. Reading, Massachusetts: Addison-Wesley, 1967.

Gordon, E. W. Utilizing Available Information from Compensatory Education Surveys (Final Report). Washington, D. C.: U.S. Office of Education, 1971.

Greene, D. S. "A 'New' School Idea: No Pass, No Promotion," The Observer. November, 1967, p. 1.

Kidder, S. Quantity and Quality of Instruction. Paper presented to the American Educational Research Association, 1975.

Michigan Department of Education. Report of the 1974-75 Michigan Cost-Effectiveness Study. Washington, D. C.: Capital Publications, 1976.

O'Reilly, R. The Contributions of Quantity and Quality of Instruction to Reading Programs. Paper presented to the American Educational Research Association, 1975.

Penny, R. E. Presenting Behaviors and Teaching Success. Unpublished doctoral dissertation, Stanford University, 1969.

St. John, N. H. School Desegregation: Outcomes for Children. New York: John Wiley & Sons, 1975.

REFERENCES (Cont'd)

Soar, R. S. An Integrative Approach to Classroom Learning. Philadelphia, Pennsylvania: Temple University, 1966.

Trisman, A.; Waller, M. I.; and Wilder, G. A Descriptive and Analytic Study of Compensatory Reading Programs: Final Report (Volume I). Princeton, New Jersey: Educational Testing Service, 1975.